

# Pengecaman Sebutan Digit Menggunakan Rangkaian Neural : Satu Kajian Terhadap Carian Bilangan Nod.Tersembunyi dan Parameter Pembelajaran

Md Sah Hj Salam\*, Dzulkifli Mohamad\*, Sheikh Husain Sheikh Salleh\*\*

\* Fakulti Sains Komputer Sistem Maklumat U.T.M.

\*\* Fakulti Kejuruteraan Elektrik U.T.M .

E-mel : sah@fsksm.utm.my

## ABSTRAK

Kajian ini adalah berkenaan dengan pemilihan bilangan nod tersembunyi dan parameter pembelajaran yang bersesuaian di dalam proses pengecaman sebutan digit bahasa Melayu dengan menggunakan rangkaian neural. Pendekatan yang digunakan di dalam melatih rangkaian neural *Feed Forward Multilayer Perceptron* (FFMLP) adalah dengan menggunakan teknik pembelajaran rambatan balik. Penganalisan suara di dalam kajian ini menggunakan kaedah *Linear Predictive Coding* (LPC) dan *Log Area Ratio* (LAR) mewakili isyarat ucapan bagi setiap 20ms. Operasi pembelajaran di dalam rangkaian neural banyak dipengaruhi oleh pemilihan parameter momentum,  $\beta$  dan kadar pembelajaran,  $\alpha$  dan bilangan nod tersembunyi. Ketidaksesuaian nilai parameter dan bilangan nod di aras tersembunyi akan menyebabkan rangkaian memberi jawapan yang mencapah daripada matlamat sebenar.

kata kunci : Sistem Pengecaman Suara, Rangkaian Neural, Nod Tersembunyi, Momentum, Kadar Belajar.

## ABSTRACT

This research is about choosing suitable number of nodes in hidden layer and learning parameters in the proses of Malay number speech recognition system using neural network. The learning approach used in training *Feed Forward Multilayer Perceptron* (FFMLP) is back propagation. Speech for the study are analysed using *Linear Predictive Coding* (LPC) and *Log Area Ratio* (LAR) to represent speech signal for every 20ms. The neural network learning operation are greatly influenced by the parameters ie. momentum, learning rate and number of hidden nodes choosen. Choosing unsuitable parameters will make the neural network system give an answer divert from the desired one.

## 1.0 Pengenalan

Pengecaman suara merupakan satu bidang kajian yang penting. Penguasaan terhadap teknologi ini akan memberi perubahan yang besar terhadap ekonomi dan cara hidup manusia. Jika sekiranya interaksi manusia dengan mesin selama ini menggunakan alat seperti tetikus dan papan kekunci yang memerlukan pergerakan tangan, teknologi pengecaman suara boleh merubah

## Pengecaman Sebutan Digit Menggunakan Rangkaian Neural : Satu Kajian Terhadap Carian Bilangan Nod.Tersembunyi dan Parameter Pembelajaran

Md Sah Hj Salam\*, Dzulkifli Mohamad\*, Sheikh Husain Sheikh Salleh\*\*

\* Fakulti Sains Komputer Sistem Maklumat U.T.M.

\*\* Fakulti Kejuruteraan Elektrik U.T.M .

E-mel : sah@fsksm.utm.my

### ABSTRAK

Kajian ini adalah berkenaan dengan pemilihan bilangan nod tersembunyi dan parameter pembelajaran yang bersesuaian di dalam proses pengecaman sebutan digit bahasa Melayu dengan menggunakan rangkaian neural. Pendekatan yang digunakan di dalam melatih rangkaian neural *Feed Forward Multilayer Perceptron* (FFMLP) adalah dengan menggunakan teknik pembelajaran rambatan balik. Penganalisan suara di dalam kajian ini menggunakan kaedah *Linear Predictive Coding* (LPC) dan *Log Area Ratio* (LAR) mewakili isyarat ucapan bagi setiap 20ms. Operasi pembelajaran di dalam rangkaian neural banyak dipengaruhi oleh pemilihan parameter momentum,  $\beta$  dan kadar pembelajaran,  $\alpha$  dan bilangan nod tersembunyi. Ketidaksesuaian nilai parameter dan bilangan nod di aras tersembunyi akan menyebabkan rangkaian memberi jawapan yang mencapah daripada matlamat sebenar.

kata kunci : Sistem Pengecaman Suara, Rangkaian Neural, Nod Tersembunyi, Momentum, Kadar Belajar.

### ABSTRACT

This research is about choosing suitable number of nodes in hidden layer and learning parameters in the proses of Malay number speech recognition system using neural network. The learning approach used in training *Feed Forward Multilayer Perceptron* (FFMLP) is back propagation. Speech for the study are analysed using *Linear Predictive Coding* (LPC) and *Log Area Ratio* (LAR) to represent speech signal for every 20ms. The neural network learning operation are greatly influenced by the parameters ie. momentum , learning rate and number of hidden nodes choosen. Choosing unsuitable parameters will make the neural network system give an answer divert from the desired one.

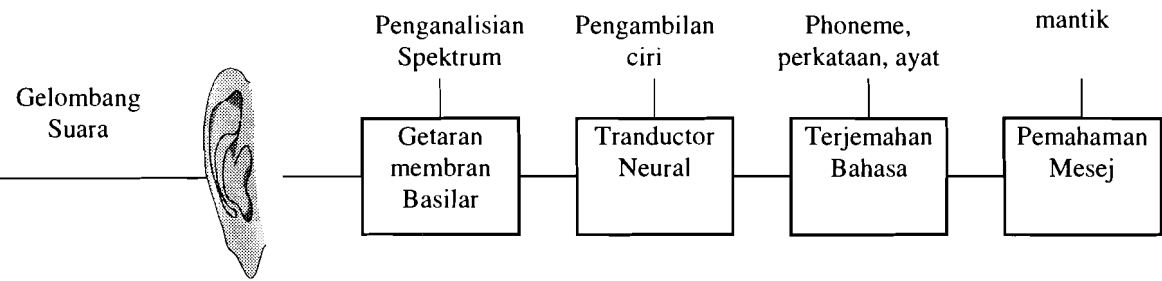
## 1.0 Pengenalan

Pengecaman suara merupakan satu bidang kajian yang penting. Penguasaan terhadap teknologi ini akan memberi perubahan yang besar terhadap ekonomi dan cara hidup manusia. Jika sekiranya interaksi manusia dengan mesin selama ini menggunakan alat seperti tetikus dan papan kekunci yang memerlukan pergerakan tangan, teknologi pengecaman suara boleh merubah

keadaan ini dengan membolehkan interaksi mesin-manusia menggunakan perantaraan suara yang lebih pantas, mudah dan selesa.

Teknologi suara ini merupakan teknologi yang telah dipelopori sejak tahun 1950'an [1]. Walaubagaimanapun, sistem yang berkeupayaan seperti manusia dalam mengenal ucapan manusia masih tidak diperolehi hingga kini. Tetapi teknologi ini telah mencapai tahap yang telah diaplikasikan pada bidang-bidang tertentu seperti urusan penerbangan, pembuatan kereta, permintaan pengambilan pakej hantaran di pejabat pos [2] dan juga telah dikomersialkan dalam berbagai bentuk.

Menurut Pierce, kerja pengecaman suara oleh mesin adalah amat sukar kerana ia memerlukan keupayaan mesin untuk mempunyai pengetahuan yang mendalam dalam ilmu lingustik, bahasa perbualan biasa dan pengalaman perbualan manusia [3]. Pendekatan yang popular dalam kajian teknologi suara ini ialah menggunakan kepintaran buatan. Ia menggabungkan bidang pengecaman pola dengan keupayaan mesin melihat, menganalisa, belajar dan membuat keputusan seumpama keupayaan seorang manusia. Kajian yang dibuat ini menggunakan rangkaian neural untuk pengecaman digit bahasa Melayu. Kertas kerja ini akan melaporkan pencarian parameter pembelajaran dan bilangan nod tersembunyi yang sesuai bagi sistem rangkaian neural yang dibina.



Rajah 1 : Pengecaman suara oleh manusia

2.0 Model Sistem Pengecaman Suara

Proses pengecaman ucapan oleh manusia bermula dengan menerima suara oleh telinga. Suara itu akan melalui membran basilar (gegendang) yang terletak pada bahagian telinga dalam yang mana gelombang suara akan dianalisa dan menghasilkan isyarat spektrum. Isyarat spektrum ini akan melalui tranduktor neural yang akan menukarkan isyarat tersebut kepada



“kosong” hingga “sembilan” sebanyak 20 kali pada waktu yang berbeza bagi mendapat variasi dalam sebutan. Mereka diminta menyebut dengan jelas tanpa menggunakan loghat. Suara ini disampel, didigit dan ditapis menggunakan perisian *GoldWave*. Frekuensi suara manusia berada dalam lingkungan 0 - 5KHz [4] oleh itu saiz sampel yang dipilih ialah 8KHz iaitu lebih kurang dua kali ganda frekuensi sebutan manusia.

Suara manusia adalah dalam bentuk isyarat selanjar (*continuous signal*). Walaubagaimanapun pada satu jangkamasa yang pendek ia dianggarkan sebagai statik [5][6]. Oleh itu bagi kajian ini setiap 20msaat suara akan dianalisa menggunakan penganalisa spektrum LPC<sup>1</sup>, dan sebanyak 12 LAR akan diambil untuk mewakili ucapan pada jangkamasa tersebut yang akan mewakili satu bingkai. Jadi, jika sebutan “kosong” mengambil masa 1.4 saat, maka jumlah bingkai yang didapati bagi mewakili sebutan tersebut adalah 70 bingkai (1400/20 msaat).

Kebiasaannya ucapan manusia tidak sama apabila disebut pada ketika yang berbeza sama ada dari segi bunyi, bentuk pola dan juga jangkamasa sebutan, walaupun menyebut perkataan yang sama oleh pengucap yang serupa. Oleh itu bilangan bingkai yang didapati pada setiap kali sebutan akan berbeza-beza dan perlu dinormalkan sebelum disuapkan kepada rangkaian neural bagi tujuan latihan dan penilaian.

Penormalan bingkai dibuat dengan mengambil kira 3 kemungkinan yang pasti berlaku selepas proses pengekstrakan ciri. 3 kemungkinan tersebut ialah

- a) Bilangan bingkai lebih besar daripada bilangan bingkai input pada rangkaian
- b) Bilangan bingkai lebih kecil daripada bilangan bingkai input pada rangkaian
- c) Bilangan bingkai sama dengan bilangan bingkai input pada rangkaian

Berdasarkan pada kemungkinan ini, proses penormalan yang dibuat seperti pada algoritma di bawah ini. Andaikan  $M = \{M_1, M_2, \dots, M_k\}$  ialah vektor ciri bagi mewakili sebutan sebelum penormalan dan  $N = \{N_1, N_2, \dots, N_j\}$  ialah vektor ciri sebutan selepas penormalan. Proses penormalan adalah seperti berikut :

```

If (Bingkai (M) == Bingkai (N) )
{ for (int k = 0; k < M; k++)           // salin semula saja
    CopySpektrum(M[k],N[k]);
}
If (Bingkai (M) > Bingkai (N))
{ ratio = Bingkai (M) / Bingkai (N) ;
  for (int k = 0; k < M; k++)           // salin mengikut ratio dan

```

<sup>1</sup> untuk maklumat lanjut bagi LPC sila rujuk Parson, Thomas W. *Voice and Speech Processing*, McGraw Hill, New York, 1996. Bab 6 (ms.137 - 167)

```

CopySpektrum (M[k * ratio], N[k]); // buang ulangan spektrum
}
if (Bingkai (M) < Bingkai( N))
{ ratio = Bingkai (N) / Bingkai (M);
  for (int k = 0; k < N; k++) // salin mengikut ratio
    CopySpektrum (M[k], N[k * ratio] );
  for ( int j = 0; j < N; j++) // isi slot yang kosong
    while (N[j] = empty) // dengan spektrum yang
      FillEmpty (N[j],N[j++]; // bersebelahan.
}

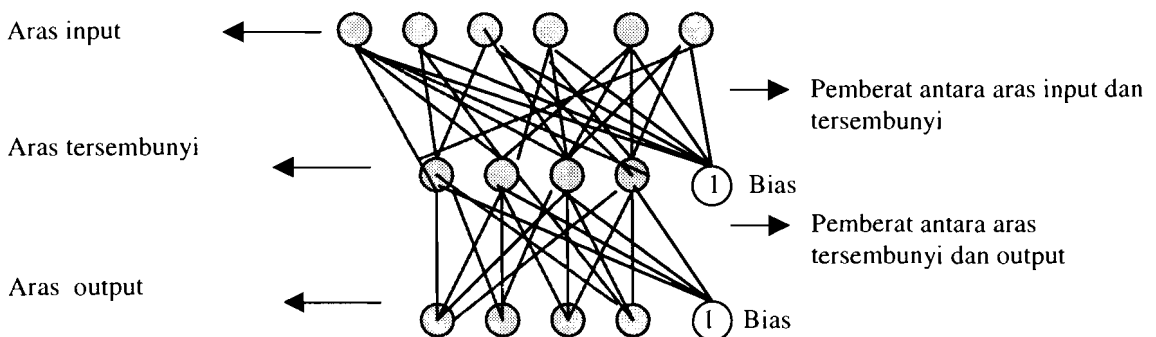
```

Pola yang dihasilkan selepas penormalan didapati tidak banyak berubah dengan pola sebelum penormalan. Lampiran 1 menunjukkan contoh hasil pola sebutan selepas penormalan kepada  $\text{Bingkai}(N) = 30$ . Graf ini diplot menggunakan perisian *excel*.

Setiap sebutan oleh pengucap akan melalui proses di atas. Jumlah semua pola yang ada ialah 800 (20 kali sebutan x 10 nombor x 4 pengucap). Sebahagian dari pola ini akan digunakan sebagai data latihan dan sebahagian lagi digunakan untuk data ujian.

#### 4.0 Rangkaian Neural

Topologi rangkaian neural yang digunakan ialah FFMLP. Banyak kajian yang menunjukkan bahawa penggunaan satu aras tersembunyi merupakan pengklasifikasi yang universal [7]. Satu keputusan ujikaji mendapati bahawa penggunaan dua aras tersembunyi tidak memberi keputusan yang lebih baik dari satu aras tersembunyi [8]. Oleh itu ujikaji ini menggunakan 3 aras rangkaian iaitu satu aras tersembunyi, satu aras input dan satu aras output. (Rajah 3).



Rajah 3 : Topologi Rangkaian Neural

Vektor ciri yang mewakili pola sebutan yang dihasilkan dalam proses penormalan akan diletakkan pada nod aras input. Nilai-nilai pada nod input ini akan dihantar ke aras tersembunyi dan kemudian ke aras output bersama-sama pemberat pada setiap hubungan antara nod. Setiap hubungan antara nod berfungsi dengan mengira jumlah bersama pemberat melalui fungsi tidak linear seperti fungsi sigmoid atau hyperbolic tangen. Ini boleh ditunjukkan seperti di bawah.

$$Out_i = f(net_i) = f\left(\sum_j W_{ij} Out_j + \theta_i\right) \quad (1)$$

$Out_i$  merupakan nilai bagi nod ke  $i$  pada aras yang hendak dihitung.  $Out_j$  pula ialah nilai nod ke  $j$  pada aras sebelum aras yang hendak dihitung.  $W_{ij}$  pula ialah pemberat pada hubungan aras di antara aras nod  $Out_i$  dan  $Out_j$ .  $\theta_i$  merupakan nilai bias pada aras yang mengalami proses pengiraan. Ia memainkan peranan dalam mempercepatkan pembelajaran rangkaian [9].

#### 4.1 Latihan dan Pembelajaran Rangkaian

Rangkaian neural yang dibina ini dilatih menggunakan skema rambatan balik latihan berpemantauan. Skema ini menggunakan ralat puncakuasa dua sebagai pengukur keberkesanan rangkaian belajar. Ralat puncakuasa dua diungkap seperti berikut

$$E = \frac{1}{2} \sum_{j=1}^N (e_j)^2 \quad (2)$$

yang mana  $e_j = (t_j - o_j)$  dan  $N$  ialah bilangan input,  $t$  ialah sasaran sebenar dan  $o$  ialah output dari rangkaian. Sebagai contoh rangkaian mempunyai 4 nod output yang mewakili nombor “sembilan” sebagai  $t[4] = \{0.1, 0.0, 0.0, 0.1\}$  manakala output dari rangkaian ialah  $o[4] = \{0.09, 0.05, 0.05, 0.09\}$ . Oleh itu pengiraan ralat pada input pertama ini,  $j = 1$  sebelum dijumlahkan keseluruhan input adalah seperti berikut.

$$\begin{aligned} e1 &= [ (t[0] - o[0])^2 + (t[1] - o[1])^2 + (t[2] - o[2])^2 + (t[3] - o[3])^2 ] \\ &= 0.0001 + 0.0025 + 0.0025 + 0.0001 \\ &= 0.0052 \end{aligned}$$

Setiap jumlah  $e_j$  seperti di atas akan memberi kesan pada ralat pembelajaran terhadap rangkaian. Jika ralat antara sasaran dengan output dari rangkaian adalah besar, maka jumlah ralat akan didapati besar.

Pembelajaran dalam skema rambatan balik dibuat dengan merendahkan nilai ralat rangkaian melalui pengemaskinian pemberat menggunakan kaedah *steepest descend*<sup>2</sup>. Ringkasan kaedah ini boleh diungkapkan sebagai ,

$$\Delta w_{ji} = \eta \delta_{pj} O_{pi} \quad (3)$$

dengan  $\eta$  adalah kadar belajar,  $\delta_{pj}$  adalah ralat isyarat pada nod  $j$  pada aras  $L$  dan  $O_{pi}$  adalah output bagi nod  $i$  pada aras  $L-1$ . Nilai  $\delta_{pj}$  dikira sebagai

$$\delta_{pj} = (t_{pj} - O_{pj}) O_{pj} (1 - O_{pj}) \text{ jika nod output} \quad (4)$$

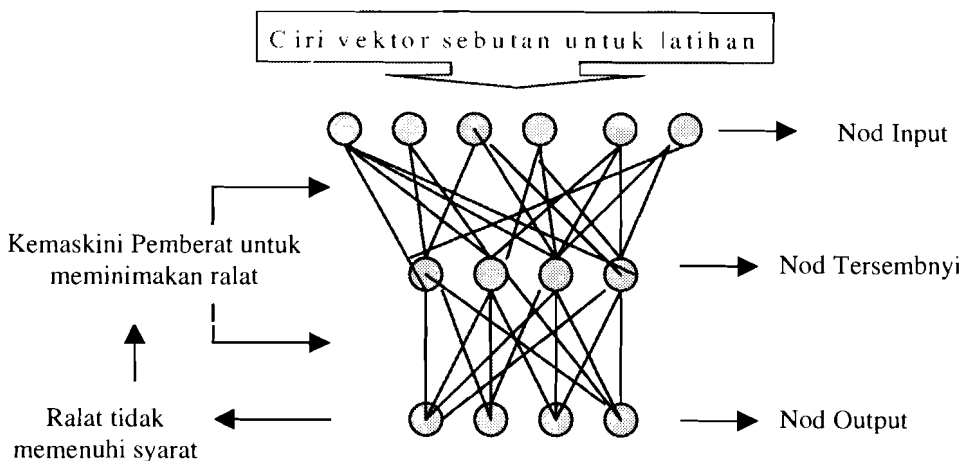
dan

$$\delta_{pj} = O_{pj} (1 - O_{pj}) \sum_k \delta_{pk} w_{kj} \text{ jika nod tersembunyi} \quad (5)$$

Dari persamaan (3), nilai kadar belajar,  $\eta$  berfungsi sebagai penentu saiz kecerunan persamaan iaitu berapa lama masa pembelajaran akan berlaku. Nilai  $\eta$  perlu bersesuaian agar penumpuan cepat tetapi tidak terlalu besar kerana di khuatiri akan menyebabkan pembelajaran menjadi tidak stabil. Bagi membantu penumpuan yang lebih cepat nilai momentum,  $\alpha$  ditambah pada persamaan (3) seperti berikut

$$\Delta w_{ji}(n+1) = \eta (\delta_{pj} O_{pi}) + \alpha \Delta w_{ji}(n) \quad (6)$$

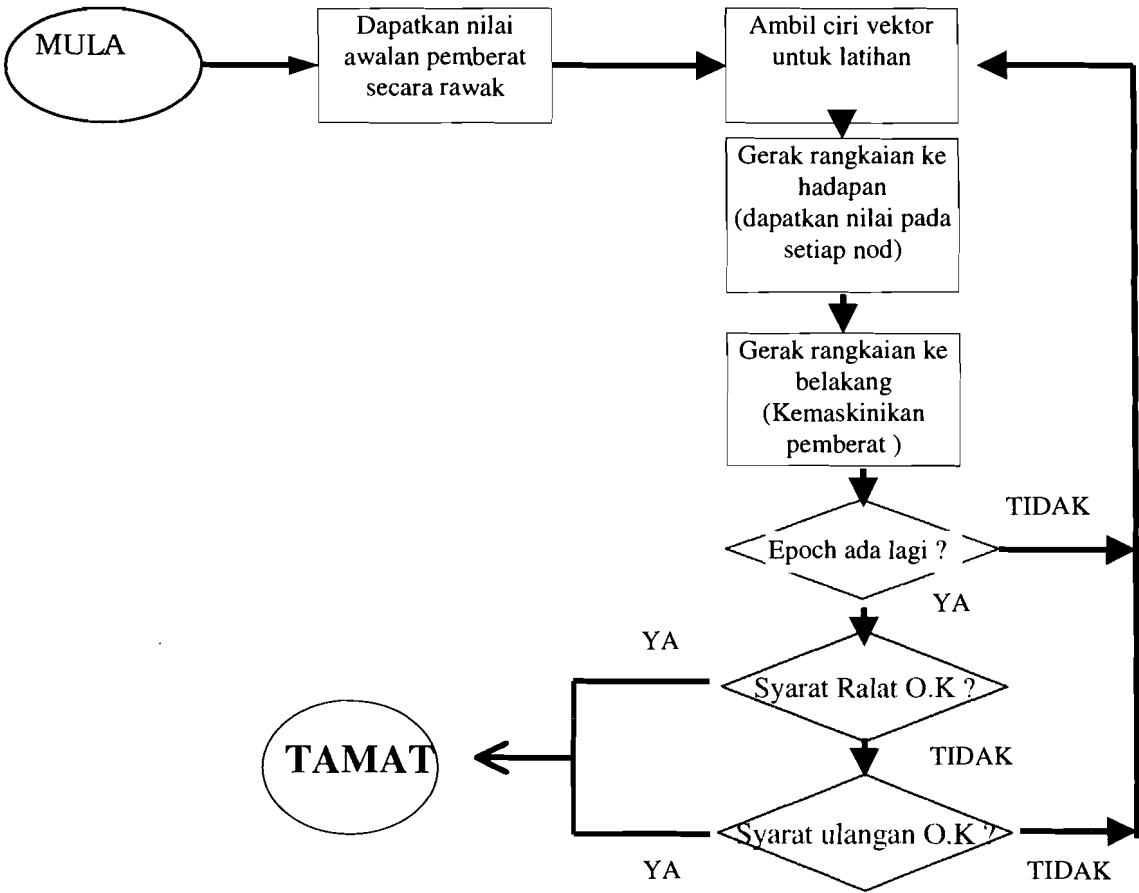
Nilai momentum ini mengambil kira perubahan pada pemberat sebelumnya. Proses latihan rangkaian adalah seperti pada Rajah 4 dan carta alir bagi algoritma rambatan balik seperti pada Rajah 5.



Rajah 4 : Proses Latihan

<sup>2</sup> sila rujuk A.S Pandya, R.B Macy *Pattern Recognition with Neural Network*, CRC Press, Florida 1996. (ms. 88 - 93) .





Rajah 5 : Algoritma Rambatan Balik

4.2 Perwakilan Nombor Sebutan Pada Nod Output.

Rangkaian ini dilatih secara berpemantauan iaitu nilai sasaran digunakan untuk melihat sejauh mana ralat dikurangkan. Oleh itu nod output perlu dapat mewakili nombor bagi sebutan “kosong” hingga “sembilan” dengan baik. Rangkaian yang dibina menggunakan empat nod dalam mewakili nombor bagi sebutan “kosong” hingga “sembilan”. Nod-nod ini akan diklasifikasikan sebagai nombor sebutan dengan memberi nilai “on” dan “off” pada nod tersebut dalam bentuk penduaan. Jadual 1 menunjukkan perwakilan nombor pada nod output. Nilai 0.0 menandakan nod pada nilai tersebut “off” manakala nilai 0.1 menandakan nod pada nilai tersebut “on” dengan nilai 0.0 dan 0.1 dipilih secara cuba-jaya.

Jadual 1 : Perwakilan nombor pada nod-output (nilai sasaran)

	0	1	2	3	4	5	6	7	8	9
O [0]	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.1
O [1]	0.0	0.0	0.1	0.1	0.0	0.0	0.1	0.1	0.0	0.0
O [2]	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.0	0.0
O [3]	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1

Setelah proses latihan dibuat, nilai nod output akan berubah bergantung pada pembelajaran rangkaian. Contoh nod output selepas latihan bagi nombor “kosong” hingga “sembilan” adalah seperti dibawah.

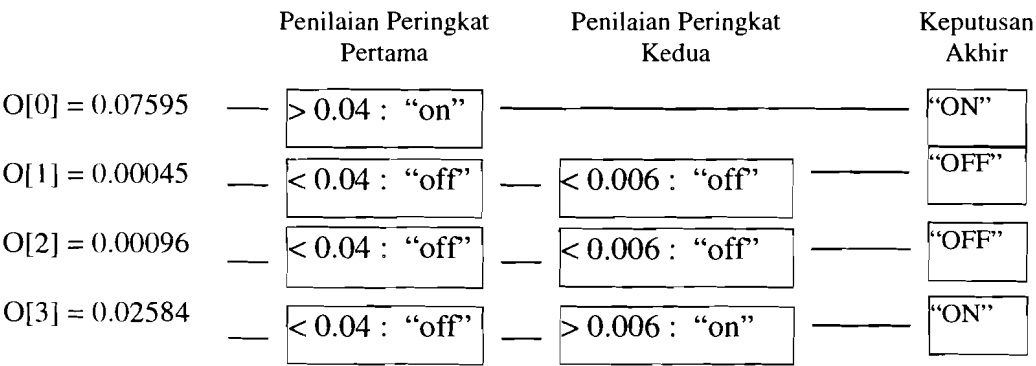
Jadual 2: Contoh nilai nod output selepas latihan (nilai dapatan)

	0	1	2	3	4	5	6	7	8	9
O [0]	0.00012	0.10219	0.00046	0.12161	0.00064	0.10786	0.00308	0.12879	0.00008	0.09872
O [1]	0.00027	0.00010	0.04808	0.06447	0.00022	0.00017	0.08423	0.10467	0.00002	0.00026
O [2]	0.00006	0.00023	0.00036	0.00034	0.04727	0.12633	0.04007	0.11871	0.00001	0.00155
O [3]	0.00008	0.00006	0.00015	0.00032	0.00019	0.00002	0.00070	0.00021	0.22652	0.04948

Boleh diperhatikan pada jadual 2, bahawa output rangkaian tidak mempunyai nilai yang sama dengan nilai sasarannya iaitu 0.0 dan 0.1. Selepas proses pembelajaran nilai yang menghampiri 0.1 akan dikira “on”, manakala nilai yang menghampiri 0.0 akan dikira “off”. Permasalahan yang timbul untuk pengkelasan ialah memilih nilai ambangan (*threshold*) yang sesuai.

### 4.3 Pengkelasan Nombor Sebutan

Pengkelasan nombor sebutan dibuat dengan meletakkan dua nilai ambangan bagi menilai satu-satu nod sama ada ia diklasifikasikan sebagai “on” atau “off”. Dua nilai tersebut ialah 0.04 dan 0.006 yang mana nilai ini dipilih secara cuba-jaya. Teknik pengklasifikasian nombor dibuat seperti contoh dibawah.



Rajah 6: Cara pengelasan

Setiap nilai pada nod output akan diperiksa dengan nilai ambangan 0.04. Jika terdapat nilai yang lebih besar, maka nod tersebut akan dilabelkan sebagai "on". Manakala nilai nod lain akan diperiksa dengan nilai ambangan kedua dan jika terdapat nilai yang lebih besar dari nilai ini maka nod tersebut juga akan di label "on". Langkah seterusnya adalah memetakan nod output dengan nombor secara penduaan.

5.0 Pencarian Parameter Rangkaian dan Nod Tersembunyi

Terdapat beberapa faktor yang memberi kesan kepada pembelajaran rangkaian. Antaranya ialah faktor pemilihan parameter dan struktur rangkaian. Sebagaimana diterangkan pada bahagian 4.1, kadar pembelajaran dan momentum memainkan peranan yang penting dalam mempercepatkan proses pembelajaran rangkaian. Manakala bilangan nod tersembunyi yang sesuai akan memberi prestasi rangkaian dengan lebih baik dalam memetakan input output rangkaian [9].

Terdapat beberapa ujikaji telah dijalankan dalam memilih nilai momentum dan kadar pembelajaran. Nilai yang terbaik walaubagaimanapun tidak sama pada semua kasus. Hussain telah menyenaraikan beberapa nilai pasangan kadar pembelajaran dan momentum yang pernah digunakan dalam pengecaman suara menggunakan rangkaian neural [10] . Dalam kajian ini, lima pasangan momentum dan kadar pembelajaran telah dipilih, yang mana empat daripada pasangan ini adalah yang pernah digunakan dalam ujian pengecaman suara. Satu nilai pasangan diambil secara rawak iaitu {0.9,1.0}. Nilai-nilai pasangan tersebut ialah.

Momentum	0.5	0.5	1.0	0.9	0.1
Kadar Belajar	0.25	0.75	0.9	1.0	0.9

Manakala dalam mencari bilangan nod tersembunyi yang sesuai, beberapa cadangan telah diperkenalkan sebelum ini. Antara cadangan itu ialah bilangan nod tersembunyi,  $h = n$ ,  $h = 2n$ ,  $h = \sqrt{n,m}$  [9] dan  $h = 3n$  [11] yang mana  $n$  bilangan nod input dan  $m$  ialah bilangan nod output. Cadangan ini tidak berapa sesuai untuk di aplikasikan pada nod input yang besar seperti pada kajian ini. Oleh itu, kami menggunakan kaedah cuba jaya dengan pemilihan nilai yang kecil iaitu bermula dari 30 hingga 70 dengan tokokan antara nombor adalah 5 iaitu { 30,35, .... 65,70}.

5.1 Ujikaji Mencari Parameter Rangkaian dan Bilangan Nod Tersembunyi

Tujuan utama ujikaji ini adalah untuk mencari pasangan momentum, kadar pembelajaran dan bilangan nod tersembunyi yang sesuai untuk pembelajaran rangkaian dalam aplikasi pengecaman suara. Ujikaji ini terbahagi kepada dua bahagian iaitu mencari pasangan momentum dan kadar pembelajaran dan bahagian kedua mencari bilangan nod tersembunyi yang sesuai.

Pada bahagian pertama, struktur rangkaian, bilangan lelaran, ralat penerimaan, nilai awalan pemberat di tetapkan seperti pada Jadual 3. Manakala dalam mencari bilangan nod tersembunyi yang sesuai, parameter kawalan di tetapkan seperti pada Jadual 4. Setiap parameter yang hendak dicari akan melalui proses latihan. Proses latihan rangkaian akan ditamatkan sama ada ralat penerimaan atau bilangan lelaran dicapai.

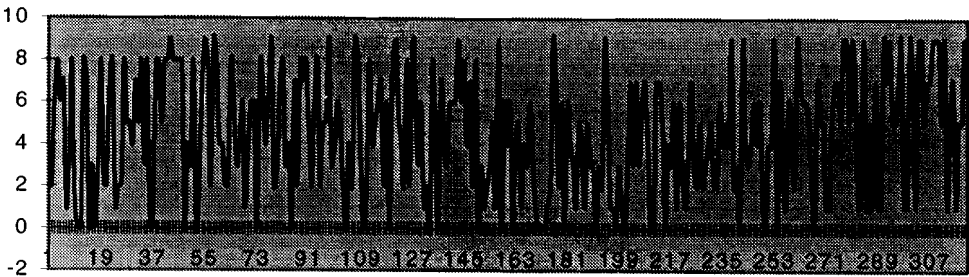
Proses latihan atau pembelajaran dibuat dengan data input sebanyak 320 sebutan dengan setiap nombor akan diperkenalkan pada rangkaian sebanyak 32 kali. Data input ini dimasukkan secara rawak, iaitu susunan nombor “kosong” hingga “sembilan” dimasukkan secara jujukan seperti pada Rajah 7. Tujuan susunan ini dibuat begitu adalah supaya rangkaian dapat membuat *generalization* terhadap pola dengan lebih baik.

Jadual 3 : Parameter Kawalan

Ralat Penerimaan : 0.0001	Struktur Rangkaian : 360 : 30: 4
Bil. Lelaran : 6000	Nilai Awalan Pemberat : [-3,3]

Jadual 4: Parameter Kawalan -

Ralat Penerimaan : 0.0001	Struktur Rangkaian : 360 : H : 4
Bil. Lelaran : 6000	Nilai Awalan Pemberat : [-3,3]
Pasangan Momentum dan Kadar Pembelajaran : **	



Rajah 7 : Susunan Data Masukan Dalam Latihan.

Selepas tamat sesi latihan, nilai-nilai pemberat yang menghubungkan setiap nod pada rangkaian akan disimpan bagi proses pengujian.

5.2 Pengujian

Proses pengujian dibuat dengan 320 pola nombor dengan setiap nombor dinilai sebanyak 32 kali. 320 pola ini merupakan pola yang belum pernah di perkenalkan pada rangkaian sebelum ini. Pengecaman dibuat secara rawak untuk memastikan bahawa rangkaian tidak menghafal pola tetapi mampu membuat *generalization*.

Pengiraan ketepatan pengecaman dibuat dengan mengira bilangan pola yang tepat dicam dibahagi dengan jumlah bilangan pola yang diuji seperti berikut,

$$\text{Ketepatan pengecaman} = \frac{\text{Bil. Pola yang tepat dicam}}{\text{Jumlah Bil. Pola yang diuji}} \times 100$$

\*\* nilai pasangan momentum dan kadar pembelajaran terbaik yang di capai pada ujikaji pertama

Selain dari melihat ketepatan pengecaman, prestasi pembelajaran rangkaian juga diambil kira dalam menilai parameter yang sesuai. Pembelajaran dilihat melalui graf ralat lawan lelaran bagi setiap set latihan.

6.0 Keputusan

6.1 Keputusan Bagi Pencarian Momentum dan Kadar Belajar.

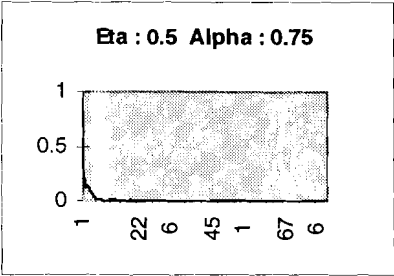
Keputusan mencari nilai pasangan momentum dan kadar belajar adalah seperti pada jadual 5. Rajah 8 hingga 12 merupakan graf ralat lawan lelaran bagi pembelajaran rangkaian untuk setiap set latihan.

Kadar Belajar (Eta)	0.25	0.5	1.0	0.9	0.1
Momentum (alpha)	0.5	0.75	0.9	1.0	0.9
Kadar Pengecaman	90%	92%	20%	<hang>	93%
Ralat Mula	1.38584	0.96896	2.26425	2.405	1.3676
Ralat Akhir	0.0003	0.00021	1.76007	2.405	0.00021

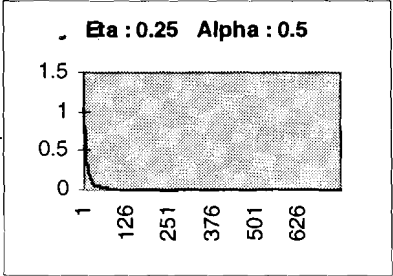
Jadual 5 : Keputusan Bagi Pasangan Momentum dan Kadar Belajar

Pada Jadual 5, tahap pengecaman bagi pasangan momentum dan kadar pembelajaran {0.25,0.5}, {0.5,0.75} dan {0.1,0.9} mencapai ketepatan melebihi 90% manakala pada pasangan {1.0,0.9} hanya mencapai ketepatan 20% dengan kadar pembelajaran yang sangat perlahan (Rajah 10). Pasangan {0.9,0.1} tidak dapat belajar langsung (Rajah 11) dimana tiada penurunan ralat setelah tamat sesi latihan dan komputer *hang* apabila pengujian dibuat.

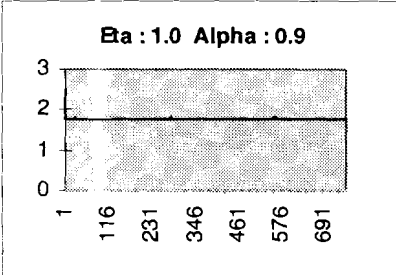
Kesemua pasangan momentum dan kadar pembelajaran diatas tidak mencapai ralat penerimaan 0.0001 dengan sesi latihan tamat setelah lelaran mencapai 6000. Walaubagaimanapun, boleh diperhatikan bahawa set pasangan yang mempunyai ralat terendah ialah {0.5,0.75} dan {0.1,0.9} iaitu 0.00021. Walaupun set pasangan {0.1,0.9} mencapai pengecaman lebih tinggi daripada pasangan {0.5,0.75}, pasangan {0.5,0.75} akan digunakan untuk sesi latihan mencari bilangan nod tersembunyi kerana ujikaji dengan struktur rangkaian yang berbeza menunjukkan ralatnya menurun dengan lebih baik.



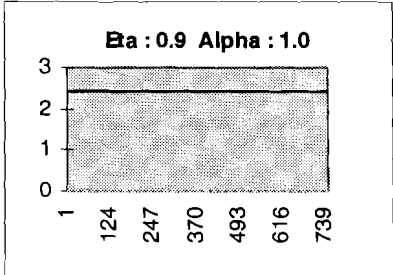
Rajah 8



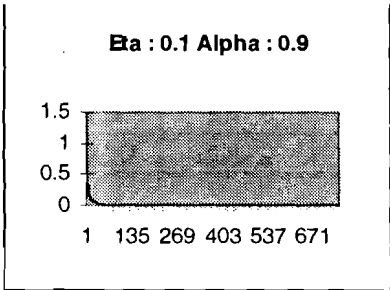
Rajah 9



Rajah 10



Rajah 11



Rajah 12

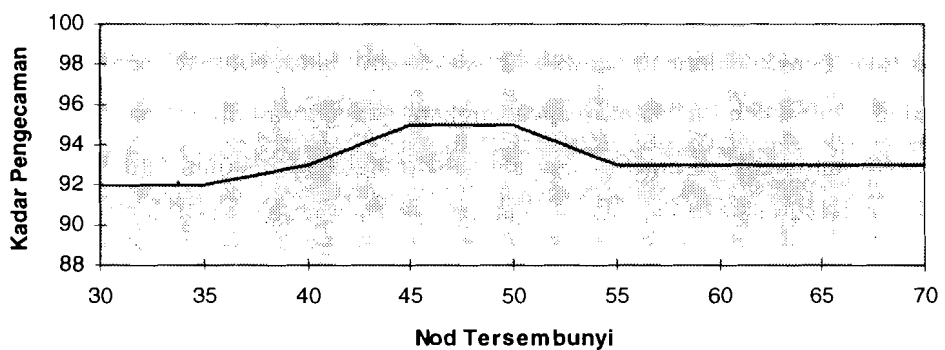
6.2 Keputusan Bagi Pencarian Bilangan Nod Tersembunyi

Keputusan bagi mencari bilangan nod tersembunyi yang sesuai adalah seperti jadual 6 di bawah. Rajah 13 menunjukkan keputusan dalam bentuk graf dan lampiran 2 menunjukkan pembelajaran untuk setiap set bagi latihan menggunakan bilangan nod dari 30 hingga 70.

Bilangan Nod Tersembunyi	Ketepatan Pengecaman	Ralat Mula	Ralat Akhir
30	92 %	0.96896	0.00021
35	92 %	0.77195	0.00026
40	93 %	0.84084	0.00093
45	95 %	0.84922	0.00010
50	95 %	0.81826	0.00029
55	93 %	0.84016	0.00053
60	93 %	1.27723	0.00547
65	93 %	0.75126	0.00475
70	93 %	1.07863	0.00086

Jadual 6 : Keputusan Ujikaji Bilangan Nod Tersembunyi

Pada Jadual 6, tahap pengecaman bagi kesemua bilangan nod tersembunyi yang dicuba melebihi 90%. Pada nilai bilangan nod tersembunyi 30, pengecaman mencapai 92% dan ia mula meningkat pada bilangan nod tersembunyi 40. Pengecaman maksima didapati pada bilangan nod tersembunyi 45 dan 50. Pengecaman mula menurun selepas itu kepada 93% hingga bilangan nod tersembunyi 70.



Rajah 13 : Graf Keputusan Pengecaman Lawan Nod Tesembunyi

## 7.0 Kesimpulan

Melalui ujikaji yang dibuat ini, boleh dilihat bahawa parameter algoritma pembelajaran dan bilangan nod tersembunyi memberi kesan pada prestasi rangkaian dalam pengecaman. Pada ujikaji mencari pasangan momentum dan kadar belajar, di dapati bahawa nilai pasangan yang



telah digunakan dalam aplikasi yang sama sebelum ini tidak semestinya terbaik untuk aplikasi yang dibuat. Ini boleh dilihat pada pasangan  $\{1.0, 0.9\}$  yang mana ia hanya mencapai ketepatan 20% dan Rajah 10 menunjukkan bahawa pembelajarannya terlalu lambat. Sekiranya pemilihan pasangan momentum dan kadar belajar ini tidak tepat seperti pada pasangan  $\{0.9, 1.0\}$  maka rangkaian tidak dapat belajar langsung. Ini boleh dilihat pada rajah 11 yang menunjukkan graf linear pada nilai ralat 2.4050 sahaja.

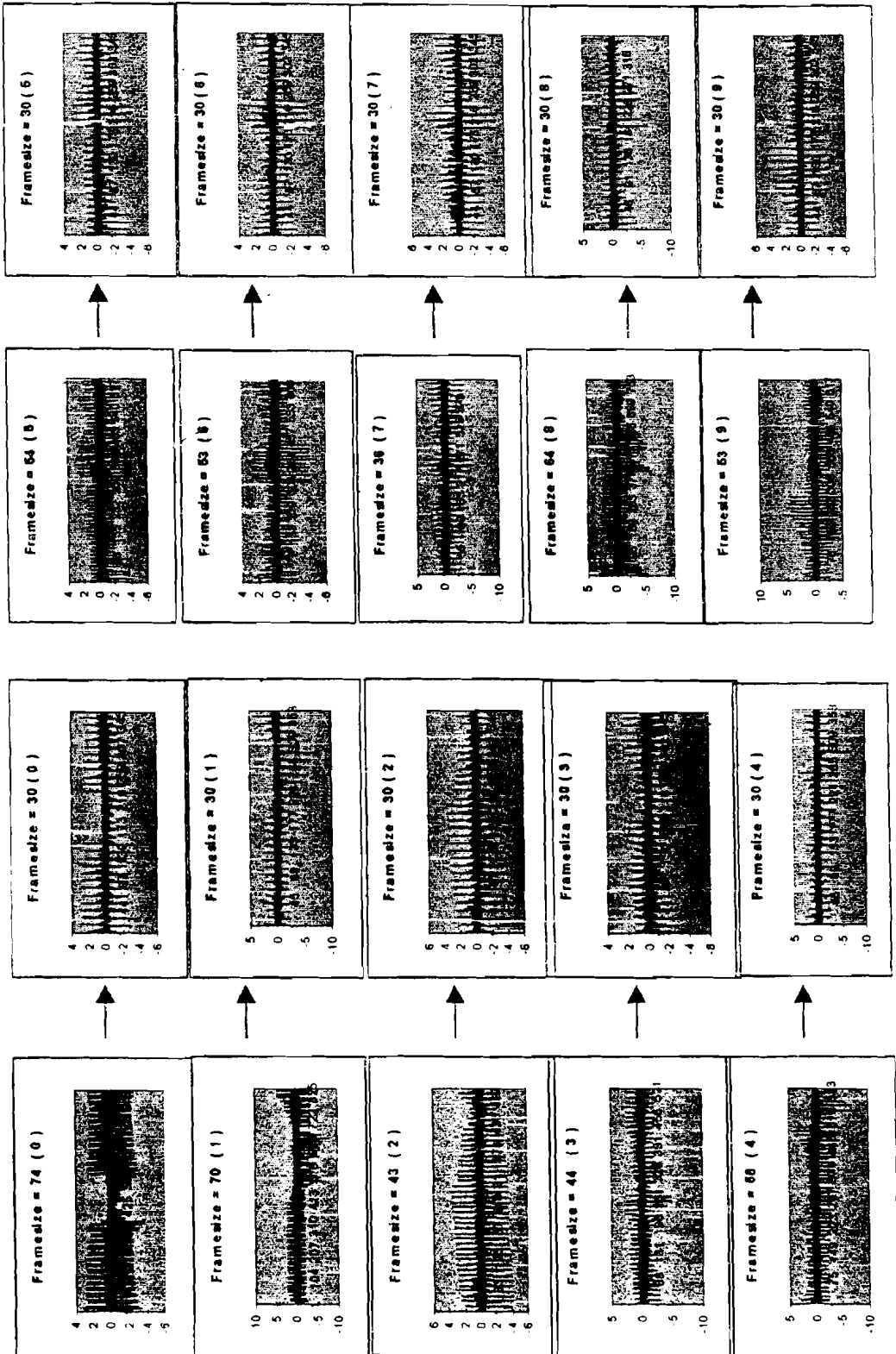
Kesan bilangan nod tersembunyi boleh dilihat pada rajah 13. Jika nilai bilangan nod ini kecil maka ketepatan maksima tidak dicapai manakala jika terlalu besar ketepatan juga akan menurun. Rajah 13 juga menunjukkan bahawa kedudukan bilangan nod yang sesuai terletak pada nilai antara 45 dan 50. Walubagaimanapun, boleh dilihat bahawa pada bilangan nod tersembunyi 45 latihan telah dapat menumpu pada ralat maksima 0.0001 dibandingkan dengan semua set latihan lain tidak menumpu pada nilai tersebut. Oleh itu bilangan nod tersembunyi 45 dipilih sebagai nilai yang terbaik pada aplikasi ini.

Semua hasil yang di dapati pada kertas kerja ini melalui kaedah cuba-jaya. Kaedah ini menjamin pemilihan parameter rangkaian dan bilangan nod tersembunyi yang baik. Walaubagaimanapun, ia memakan masa yang lama dan tidak mencuba kesemua kemungkinan jawapan. Jika sekiranya satu latihan rangkaian memerlukan masa sebanyak 8 jam maka 80 jam diperlukan dalam mencuba 10 pembolehubah yang berbeza. Walaupun pada hakikatnya mungkin nilai terbaik tidak berada dalam 10 pembolehubah tersebut iaitu mungkin memerlukan percubaan hingga 20 atau lebih. Oleh itu, kami mencadangkan penggunaan algoritma genetik (g.a) dalam mencari secara automatik parameter dan akitektur rangkaian. Kajian terdahulu menunjukkan bahawa g.a mampu mencari parameter dan akitektur rangkaian yang baik dalam masa yang lebih singkat.

## Rujukan :

1. Rabiner, L and Juang, B.H. (1993). *Fundamentals of Speech Recognition*. Englewood Cliffs, N.J : Prentice Hall. pg. 6.
2. Turban. E. (1999) *Expert System and Artificial Intelligence*. 1<sup>st</sup> ed. Republic of Singapore : MaxMillan Publishing Company. Pg. 315
3. Pierc, J.R (1969). *Whither Speech Recognition ?* JASA vol. 46. 4 pg. 1029 - 1051.
4. Saiful Akmal Khalid, Shahimi Shafiee dan Ahmad Zaki Abu Bakar (1990), *Speaker Dependent Isolated Malay Speech Recognition System*. The 3<sup>rd</sup> JSPS-VCC Seminar on Integrated Engineering: Part-B, University Malaya, Malaysia.
5. Parsons, T.W. (1987). *Voice and Speech Processing*. McGraw-Hill, New York pg. 4.
6. Sheikh Hussain Salleh, Ahmad Zuri Sha'ameri, Ahmad Idil Abd. Rahman and Zulkalnaian Mohd Yusoff. (1999). *Computer Assisted Speaker Indentification (C.A.S.I.S.)*. Jurnal Teknologi, UTM. No. 31, Dis. Pg. 29-43.
7. Pandya, A.S and Macy, R.B. (1996). *Pattern recognition with Neural Networks in c++*. CRC Press, Florida.
8. Sheikh Hussain, McInness, F.R, and Jack, M.A (1995). *Enhanced Automatic Speaker Verification Based on Combination of Hidden Markov Models and Multi Layer Perceptrons*. MICC '95, Langkawi Malaysia, 20-23 Nov.
9. Roselina Sallehuddin, Mohd Salihin Ngadimin and Siti Mariyam Shamsuddin (1999). *Penentuan Saiz dan Bilangan Nod Tersembunyi Rangkaian Neural Bagi Peramalan*. Jurnal Teknologi Maklumat, FSKSM, UTM, Jun ms. (67-78).
10. Sheikh Hussain Sheikh Salleh. (1993). *A Comparative Study of The Traditional Classifier and the Connectionist Model for Speaker Dependent Speech Recognition System*. Master Thesis, University Technology Malaysia.
11. Lipmann, R.P. (1987). *An Introduction to Computing With Neural Nets*. IEEE ASSP Magazines, April ; 4-22.

Contoh penormalan bagi sebutan “kosong” hingga “sembilan” kepada saiz frame tetap = 30



## LAMPIRAN 2

Pembelajaran rangkaian dalam bentuk graf ralat lawan lelaran bagi ujikaji pencarian bilangan nod tersembunyi bagi nilai  $h = \{30, 35, 40, 45, 50, 55, 60, 65, 70\}$ .

